DOCUMENT RESUME

ED 246 082 TM 840 353

AUTHOR Donmoyer, Robert
TITLE Choosing from Plausible Alternatives in Interpreting Qualitative Data.
PUB DATE Apr 84
NOTE 32p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Data Analysis; Epistemology; Research Methodology; Research Problems; Theories; *Validity
IDENTIFIERS *Data Interpretation; Phillips (D); *Qualitative Research; Toulmin (Stephen)

ABSTRACT

This paper addresses a variation of the traditional validity question asked of qualitative researchers. Here the question is not "How do we know the qualitative researcher's question is valid?" but rather, "How does the qualitative researcher choose from among a multitude of apparently valid or at least plausible interpretations?" As early as 1975, Popp declared that the question of how one should select from possible alternative paradigms must be dealt with by researchers. Phillips' recent attempt to salvage logical positivism, and Toulmin's belief that there is such a thing as scientific rationality and that there are ways to rationally assess the relative worth of conflicting claims both contribute to this discussion. The answer to Popp's question then serves as a basis for resolving the particular data analysis problem which provided the impetus for this paper. Finally, the paper concludes with the presentation of a series of insights about qualitative and other forms of educational research which appear to follow from the earlier analysis. (BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED246082

CHOOSING FROM PLAUSIBLE ALTERNATIVES IN INTERPRETING QUALITATIVE DATA

Robert Donmoyer

The Ohio State University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Donmoyer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the 1984 Meeting of the
American Educational Research Association
New Orleans, April, 1984

TM 840 353

The research study discussed in this paper was supported by the National Institute of Education (Grant Number G 810083). Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author and do not necessarily reflect the views of the Institute or the Department of Education.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CHOOSING FROM PLAUSIBLE ALTERNATIVES IN INTERPRETING QUALITATIVE DATA

Robert Donmoyer

The Ohio State University

This paper focuses on a data analysis problem which emerged after I had completed the final report for a study. The study examined a team of parents and school professionals employing a deliberative approach to evaluation to assess a school program (Donmoyer, 1983). Throughout the process, professionals consistently dismissed as unimportant concerns expressed by parents and also rejected parent recommendations for quite minimal changes in program procedures. The final report presented three quite different interpretations of this recurring phenomenon, although one of these interpretations was judged "most powerful," and it received a bulk of attention in the report. Later, additional interpretations were suggested by the school professionals who had been studied and by a critic who reviewed a paper based on tne final report at the 1983 meeting of the American Educational Research Association (Miles, 1983).

Interestingly those who proposed alternative interpretations did not challenge my initial interpretations. Furthermore, I believe the interpretations suggested by the critic and by research subjects are probably valid. I also believe, however, that the interpretation which was initially judged, "most powerful," continues to provide the most potent account of the

data. The question, what constitutes interpretative potency, however, is less easy to resolve and is the subject of this paper.

This paper, in other words, addresses a variation of the traditional validity question asked of qualitative researchers. Here the question is not: How do we know the qualitative researcher's interpretation is valid? Rather the question is: How does the qualitative researcher choose from among a multitude of apparently valid or at least plausible interpretations?

To a large extent, the question I am asking in this paper is simply a methodological translation of the epistemological question implicit in philosophers' discussions of the role of conceptual frameworks or paradigms in empirical inquiry. As early as 1975, for example, Popp related Kuhn's work on paradigm revolutions in natural science to educational science and declared,

> It will not do to instruct the profession to always select the framework which is best supported by evidence, for what counts as evidence is determined by some disciplinary matrix. Practitioners may believe that they have selected a paradigm on the basis of its empirical warrant, but their decision as to what is to be taken as evidence is in effect a prior commitment to some way of seeing things.(pp.34)

Popp concluded his discussion by declaring that the question of "[h]ow one should select from possible alternative paradigms is an immediate, non-aviodable problem which...must be dealt with by...researchers..." (p.39). The question with which Popp concluded his discussion will serve as a

starting point here. The answer to that epistemological question will then serve as a basis for resolving the particular data analysis problem which provided the impetus for writing this paper. Finally, the paper concludes with the presentation of a series of insights about qualitative and other forms of educational research which appear to follow from the analysis contained in the paper's first two sections.

Phillips' Reconditioned Positivism

A sensible starting point for tackling Popp's epistemological question is Phillips' (1983) recent attempt to salvage logical positivism, for, if Phillips' rescue attempt is successful, Popp's epistemological difficulty (i.e., the problem of empirical warrent) disappears and the methodological question emerging from my study (i.e., the question of how to choose among multiple valid interpretations) becomes nonsensical. It is precisely the claim that multiple valid interpretations are possible, in fact, which appears to have motivated Phillips to attempt to resurrect positivism. In his paper he rails against a rampant relativism which he perceives as having taken over the field of educational research. This relativism, Phillips argues, implies a challenge which he states succinctly: "[T]here has to be some way of validating or refuting...claims. To deny this is to do away with the concept of a mistake" (pp.11-12).

Phillips' decision to look to logical positivism in responding to this challenge is somewhat surprising. As latter parts of this discussion will indicate, resurrecting logical positivism is not the only option open to one who wishes to re-introduce the concept of a mistake to educational inquiry.

Furthermore, logical positivism is a movement which has been judged almost entirely wrong even by its once staunchest supporters (Hanfling, 1981).

Phillips, of course, does not claim to embrace the wrong aspects of positivism. Which aspects of positivism he is embracing is not clear, however. Although a major purpose of Phillip's article is to explicate the multiple meanings that have been attached to the term, positivism, and to chide those within the field of education who have used the term positivism without indicating which meanings they are talking about, Phillips, himself, is quite imprecise in his use of the term. For example, Popper and Hempel are two philosophers of science who agree on most major issues, yet Phillips calls Hempel a logical positivist but declares that Popper is "not of the positivist school" but rather "was always a staunch opponent" (p. 8). Phillip's classifications seem especially odd since he mentions Hempel in the course of discussing views of the nature of theoretical constructs, and, in fact, cites Hempel's instrumentalist views as evidence that logical positivists were not realists. What Phillips (using a classification scheme developed by Nagel (1961) calls an instrumentalist view, however, also figures prominently in Popper's critique of logical positivism.

Phillips becomes no clearer when the discussion turns to which elements of positivism are to be salvaged to serve as an antidote to the relativism of those he criticizes. The centerpiece of Phillip's position appears to be the related distinctions between meaning and truth on the one hand and knowledge and belief on the other.

Concerning the first distinction, Phillips chides contemporary critics of positivism for failing to make the distinction, and then goes on to laud the logical positivists for maintaining what he calls "a crucial distinction":

> As far as the logical positivists were concerned, for a proposition to be meaningful (i.e., for it to be a proposition at all) was for it to be testable or verifiable in principle (unless it was logically true), but to be regarded as true it had actually to have undergone successfullly such an experiential test. If an untested (but testable) proposition were accepted, the person concerned could claim to have belief but not knowledge. In other words, knowledge is "justified true belief." (p.11)

Despite his obvious admiration for the logical positivists, it is unlikely that Phillips endorses their criterion for meaning since earlier in his paper he acknowledges that the logical positivists' criterion for meaning itself rested on metaphysical (i.e., nontestable) assumptions and, hence, was, by its own test, meaningless. Phillips suggests no alternative criteria for assessing meaning, however.

The closest Phillips comes to confronting the question of criteria for assessing meaning is when he criticizes Eisner's expansive conception of knowledge. "For Eisner," Phillips writes,

> Miller's Death of a Salesman, his memory of his own mother's face, a Bach fugue, a portrait by Velasquez, and a poem by Dylan Thomas or e.e. cummings, are all statements of knowledge. But he gives no indication of how the "prior" question of the meaning of these various statements is to

> be settled (is it not possible for me to claim to understand the meaning of any of these, but nevertheless to get it wrong -- how would anyone ever know?); and he does not say how these knowledge claims are to be investigated (p.11).

Phillips' position here differs in two important ways from his own characterization of the logical positivists' position. First the sorts of questions of meaning cited as examples in Phillips' criticism of Eisner refer to questions about the meaning of phenomena not to questions about the meaningfulness of propositions. Second, Phillips implies that questions of meaning are really questions of rightness and wrongness. ("[I]s it not possible for me to claim to understand the meaning...," he writes, "but nevertheless to get it wrong..."): Hence, for Phillips, questions of meaning and questions of truth appear to be inextricably linked and interdependent and questions of meaning, rather than being "prior" questions, can, in fact, only be answered <u>after</u> the truth of a statement of meaning has been established.

Though his position has little in common with the logical positivists', Phillips' linking of questions of meaning with questions of truth seems quite sensible. Many of the most important questions within the field of education are questions of meaning of the kind exemplified in Phillips' critique of Eisner. We want to know whether the characteristics exhibited by a large number of low SES students means they are culturally disadvantaged or whether they are merely culturally different. We want to know whether a student's inability to solve an arithmetic problem means that appropriate behavior has not been reinforced by educators or whether the child has simply not yet

constructed the kind of understanding Piaget describes as conservation of number. We want to know, in other words, what are the correct meanings to attribute to empirical phenomena, and it seems perfectly sensible for Phillips to assume that the term, "correct," as it is used in the first part of this sentence, is merely a synonym for "true."

Phillips, however, fails to indicate how truth is to be assessed. His references to logical positivism suggest that he is endorsing some form of empirical test, but he qualifies such support by acknowledging that critics of positivism are "right to say that there is no critical test, (i.e., no single test) that can be used to decide between rival positions" (p. 11). Phillips says nothing, however, about what is to replace the critical experiment as a basis for empirically resolving theoretical disputes. Rather he simply asserts that in time all theoretical disputes can be settled and then attempts to justify his assertion in two ways.

One justification involves the presentation of a series of examples which appear to be designed to form a kind of argumentum ad absurdum. Phillips notes, for example, the belief that the world is flat is now a totally discredited view.

The problem with this example is that it appears to be more the exception than the rule. Especially in the social sciences, agreement about anything but the most straightforward facts is difficult to come by, and the work of Habermas and others who have written about the role political

interests play in the legitimization of knowledge suggests that even if theoretical disputes are resolved, such resolutions do not necessarily imply that agreement has emerged from empirical testing. Furthermore, even in the natural sciences, different interpretations of phenomena seem more a permanent fixture than a temporary state of affairs.

Theoretical physicists, for example, have long rejected the Newtonian, cause-effect view of the universe and replaced it with a view of the universe as a place where events occur randomly. Many engineers who employ principles of physics, however, as well as most experimental scientists engaged in basic research in natural science disciplines other than physics continue to operate from a Newtonian perspective. Even the theoretical physicists with whom I am acquainted begin to think in Newtonian terms when they leave their linear accelerators and place their feet on the accelerators of their automobiles.

The reason why many of the disputes in the social and even the physical sciences will not in time be resolved empirically the way disputes over the shape of the earth have been resolved is that many questions currently being debated are questions of meaning (i.e., the meaning of phenomena) and questions of meaning are fundamentally different from questions of fact. Indeed the latter can only be answered after the former questions have been settled, after, for example, we agree about the meaning of terms such as round, flat, and earth.

Phillips quite unintentionally dramatizes the difference between

questions of meaning and questions of fact in the process of developing his argumentum ad absurdum. He writes, "I may claim my mother had purple hair, or that I know Mozart wrote Oliver Twist, or that I know the meaning of Tchaikovsky's Fourth Symphony, but it is conceivable that I am mistaken, so there has to be some way of validating or refuting these claims." (pp. 11, 12).

Now the first two items Phillips lists, claims about the color of his mother's hair and the authorship of Oliver Twist, involve the same sort of question as the question about the earth's shape. Glib assertions that claims are validated in a manner consistent with the logical positivists' verification criterion are not problematic here, at least if we restrict our frame of reference to Western cultures. (Even the claim about the color of someone's hair becomes problematic if we include among our judges the Hanunoo people of the Philippines; Conklin (1955) reports that the Hanunoo have only four terms to describe colors and that these terms also simultaneously describe the texture of the object). Even in Western cultures, however, an appeal to the logical positivists' verification criterion becomes problematic when we turn to the third item on the list, the meaning of Tchaikovsky's Fourth Symphony.

At the very least, Tchaikovsky's Fourth will mean different things to Tchaikovsky, to a maestro conducting Tchaikovsky, to the orchestra's business manager who knows that ticket sales increase when any piece by Tchaikovsky is on a concert's program, to a philosopher who has bought a ticket to the concert which includes Tchaikovsky's Fourth, and to the same philosopher who

uses Tchaikovsky's Fourth to make a point about the relationship of meaning and truth. The meaning of a symphony, in other words, is not absolute but relative . Yet there is nothing in logical positivism or in Phillips' reformulation of logical positivism that can adequately deal with this fact.

Many of the questions encountered within the field of education and educational research, however, are questions of meaning similar to the question Phillips has posed concerning the meaning of Tchaikovsky's Fourth. The phenomenon of Black English, for example, gives rise to such questions: Is Black English to be viewed by educators as a sign of cultural difference or of cultural disadvantage? If educators listen to linguists they will probably accept the former characterization and be quite amenable to accepting language diversity in their classrooms; from the linguist's perspective, language is always in the process of changing and, therefore, the whole notion of standard English is, for most linguists, a fiction. On the other hand, business executives and even many college professors outside the discipline of linguistics will likely think of both standard English and, hence, Black English, in a different way. As a result, they will probably be less likely to tolerate Black English in schools. They will likely focus on the fact that the person who can only communicate skillfully in Black English will be at a severe disadvantage not only in the market place but also in academia (most likely including the linguistics professor's classroom). The point is, different people attribute different meanings to the term, standard English, and it seems nonsensical to assume that empirical evidence alone can discredit any of these different meanings. Indeed, what counts as empirical

evidence depends on whether one is starting from the academic perspective of the linguist or the economic perspective of the business executive.

Furthermore, sometimes disputes over meaning are characterized by rivals using totally different terms. Even when reasonable translations can be made, the translated statement is never quite equivalent to the original just as a French translation of a passage in Spanish never says quite the same thing as the original. The situation is further complicated by the fact that there are certain ideas which defy translation. There is no equivalent in behaviorists' vocabulary for the Piagetian notion of conservation of number, for example; similarly the notion of reinforcement cannot be translated into the Piagetians' specialized vocabulary.

Given this translation problem, it is difficult to imagine how disputes over meaning can be resolved in any way consistent with the logical positivists' approach to verification. Indeed, the very idea that we will someday be able to decide with empirical evidence alone whether Piagetian psychology is truer than behaviorist psychology makes no more sense than the idea that we will someday be able to determine whether French is truer than Spanish. We can certainly determine whether one language is better than another in certain contexts and for certain purposes. If we are in Spain, for example, Spanish will probably be the better language to speak in most situations. Similarly, one language may be more capable of expressing certain ideas better than another. There is nothing in Phillips' reworking of logical positivism, however, to accommodate the notion of context and

purpose in settling disputes about meaning.

Philips attempts to justify his faith in logical positivism in one additional way: an appeal to authority. Phillips notes, for example, that "the theses of relativism, incommensurability, holism of meaning, and the rest, that Kuhn's work contains (at least in embryo), have been shown to be dubious at best, and it would seem that the weight of scholarship now is flowing against Kuhn" (p. 11).

Phillips' appeal to authority is no more compelling than his attempt at constructing an argumentum ad absurdum, however. Phillips is certainly correct when he asserts that "the philosophical community at large is united behind the view that to be regarded as knowledge the belief in question has to be justified in some way." Even the supposedly relativist Kuhn has indicated that the superiority of one paradigm over a rival one can be justified by appealing to such generally accepted criteria as simplicity, scope, and accuracy. Yet even if Kuhn's criteria for justification are rated less than satisfactory, it does not follow that philosophers have once again embraced the discredited tests of meaning and truth put forth by the logical positivists. As Phillips goes on to say, "[T]he current disputes among philosophers of science essentially focus on what constitutes a satisfactory justification" (p. 11). Resurrecting logical positivism and employing its notions of meaning, truth, belief, and knowledge merely as slogans contributes nothing to resolving this current epistemological dispute and certainly provides no help in solving the data analysis problem which is the

central concern of this paper.

The Toulmin Alternative

As Phillips indicates, contemporary philosophers of science do provide alternatives to logical positivism. Toulmin (1972, 1961), for example, articulates a position which responds to Phillips' quite legitimate concern that there be some way of validating or refuting researchers' claims. Indeed, what Phillips calls "the concept of a mistake" is very much a part of Toulmin's thinking. Toulmin's thinking, however, also accommodates the problem of the relativity of meaning, a problem which, as prior discussion has shown, Phillips' reconditioned positivism does not even confront.

Toulmin stakes out a middle ground position between the position of other post-positivist philosophers of science such as Kuhn (1962), who, as noted above, endorses universal but nebulous criteria for assessing the worth of conflicting paradigms, and Feyerabend (1970), whose self-described anarchistic view of science suggests that decisions concerning the relative worth of conflicting paradigms are non-rational. Toulmin clearly believes that there is such a thing as scientific rationality and that there are ways to rationally assess the relative worth of conflicting claims. He rejects, however, the idea that scientific rationality involves the application of a priori criteria common to all sciences, even the sorts of non-operational criteria suggested by Kuhn. Rather, Toulmin emphasizes the role that purpose plays in science, and he argues that differing purposes will inevitably result in different criteria for appraising the relative adequacy of

conflicting conceptual orientations.

In emphasizing the role of purpose in science, Toulmin directly challenges the common assumption that all scientific activity can be characterized by a single unitary purpose, prediction. Toulmin makes his point by means of an analogy. He suggests that the question, "What is the purpose of science?" is similar to the question, "What is the purpose of games?" Toulmin notes that a quick response to this latter question might be that the purpose of a game is to beat one's opponent. This response, Toulmin contends, would not be satisfactory since there are many games--Toulmin cites solitaire, but one could also point to a host of games involving cooperation rather than competition--which do not involve an opponent just as there are some sciences which are primarily concerned with classification and not prediction. If terminology is made more limited--if, for example, one merely asserts that the purpose of competitive games is to beat one's opponent -- Toulmin argues that one still has offered a relatively meaningless tautology. Toulmin writes:

> Competitive games are of course competitive, and this fact determines the general <u>sorts</u> of purposes they will have. Yet in playing any particular game a man will be trying to do a large number of different things--to ace his oponent, tire him out, get him out of position, and so on--pursuing a range of different aims, any of which may contribute to his overall success. It will be his business to add to his score of points or goals in whatever way will best promote the general strategy of his activity. And we shall understand the game in question only when we understand, in outline at any rate, this whole plural range of aims and purposes which a participant in it has to pursue. (1961, p. 20)

Thus Toulmin argues that each scientific discipline just like every game is undergirded by its own particular purposes. If there are fundamental disagreements over purposes within a discipline, Toulmin notes, then that discipline simply divides into sub-disciplines (e.g. linguistics divides into psycholinguistics and sociolinguistics). Toulmin also argues that the adequacy of conflicting paradigms, (e.g., the adequacy of Newtonian and Einsteinian physics) is assessed not in terms of general critical relevant to all scientific activity but in terms of the particular purpose which characterize a particular discipline.

There is an additional aspect to Toulmin's thinking which has relevance for subsequent parts of this discussion: Toulmin explicitly blurs the distinction between science and other forms of human understanding. To be sure he does point out important distinctions between science and art, for example, and science and public policy fields. Concerning the first distinction, Toulmin notes that while there is historical continuity and coherence in the development of artistic technique, all painters or poets or novelists do not share a common purpose. On the contrary, the purpose of the artist is idiosyncratic to the artist herself. At best, therefore, the arts and various genres of literature can only be considered quasi-disciplines.

Similarly, public policy fields differ from scientific disciplines because in the former there is no unanimity with respect to purpose. Toulmin dramatizes this important difference by using an example: He notes that even if civil engineers recommend that a dam be built in a particular location,

policy makers will not necessarily heed such a recommendation, for the engineers' technical concerns will only be one of a multitude of concerns which policy makers will have to take into consideration. This need to balance multiple, often conflicting concerns and purpose leads Toulmin to characterize public policy fields as "non-disciplinary and unprofessionalizable" (1972, p. 400).

The distinction Toulmin makes between scientific disciplines, quasi-disciplines and non-disciplinary fields, however, does not challenge the claim that human rationality is fundamentally the same no matter what the context. The adequacy of human constructions -- be they scientific paradigms, artistic or literary creations, or public policy decisions -- must always be judged in terms of particular purposes. It is just that in the case of scientific disciplines, debate over the purposes themselves is rare; in artistic and literary fields debate over purpose is more likely; and in the area of public policy debate about purposes is inevitable and inevitably intertwined with more technical questions about which of several policy options serve particular purposes best.

To summarize, Toulmin articulates a conception of human rationality which generalizes to all aspects of human understanding. This conception manages to resolve difficulties not even recognized by Phillips in his attempt to resurrect logical positivism. Toulmin's position, however, also manages to accommodate a concern which Phillips does recognize: The need to retain , in Phillips' words, "the concept of a mistake." In Toulmin's

conception of human rationality, in fact, two sorts of mistakes are possible. First order mistakes occur when human creations cannot adequately measure up to the purposes against which they are compared. Second order mistakes result from the inadequacy of the purposes themselves.

This discussion now turns from the somewhat ethereal, abstract realm of epistemology to the more concrete and pragmatic domain of research methods. More specifically the focus turns to the data analysis problem which emerged out of my study of a deliberative approach to program evaluation: How does one choose from among multiple, highly plausible interpretations? I believe Toulmin's analysis can provide the basis for answering this question, for, as I have indicated, the data analysis question I am asking is simply Popp's epistemological question writ small. Before I can explicate the relationship of Toulmin's analysis to the data analysis problem emerging out of my study, however, a more detailed description of the study and its analysis problem is required.

The Study

As I indicated at the outset of this paper, the study was actually a study within a study. At the center was an evaluation study. The evaluation focused on the program of an open-space public elementary school which was still trying to adhere to the tenets of open or informal education popularized during the 1960's. The evaluation was conducted by a team consisting of four teachers and the principal from the school, five parents

(some of whom were selected because they had a history of being critical of the school's program), an administrator from the district office, and a principal from another school in the district which operated a mastery learning program, a program which was dramatically different from the program being studied. A graduate research assistant and I facilitated the team's efforts as it collaboratively engaged in a process of defining what issues the evaluation would focus on, deciding what information would be useful in resolving the defined issues, gathering information, interpreting information, and making recommendations about the program being evaluated.

These procedures represent an attempt to make operational what I have referred to as a deliberative conception of evaluation. This conception is rooted in a series of assumptions including (1) the assumption that educational issues are, at base, philosophical or ideological issues which cannot be resolved by empirical evidence alone and (2) the assumption that educational disputes can be resolved through improved communication and the expanded understanding such improved communication generates rather than through the exercise of political power.

This second assumption is based on a distinction which emerges from a review of the literature on the nature of political conflict. One conception of political conflict assumes that political disputes are a product of self-interested rational actors attemting to maximize mutually exclusive preferences. From this perspective, political disputes can only be resolved through the exercise of power. A second conception of politics, and the one

on which a deliberative approach to evaluation is based, assumes that political disputes result from different perceptions of the world. Accordingly, disputes can be resolved by means of education. Through the process of communication, those who disagree can, in principle, at least, expand their understanding of an issue and adjust their egocentric perspective and action to accommodate their opponents' perspectives. Dewey (1916) articulates this position nicely:

> To be a recipient of a communication is to have an enlarged and changed experience. Nor is the one who communicates left unaffected. Try the experiment of communicating, with fullness and accuracy, some experience to another, especially if it be somewhat complicated, and you will find your own attitude toward your experience changing...The experience had to be formulated in order to be communicated. To formulate requires getting outside of it, seeing it as another would see it...One has to assimilate, imaginatively, something of another's experience in order to tell him intelligently of one's own experience (p. 5).

A grant from the National Institute of Education made it possible to informally test the validity of the above assumptions and the viability of a deliberative approach to evaluation. A study of the study was designed around the insider/outsider notion of qualitative research used by Smith and Geoffrey (1968): the research team for this second study consisted of the two researchers who were involved in the evaluation study and a third researcher playing the outsider role. Various naturalistic methods (e.g., observation, participant observation, video and audiotaping of team meetings, interviews, and questionnaires) generated a vast amount of data to be analyzed.

The Data Analysis Problem

Initially data analysis was a relatively unproblematic (albeit a cumbersome and tiring) task. The data were unequivocal with respect to the first assumption listed above: in most instances where disagreements occurred among team members, disagreements resulted from the fact that different team members were framing their questions in different ways. (For example, while certain parents wanted to talk about the lack of curriculum continuity between grade levels and among different classes on the same grade level, teachers wanted to talk about the benefits of teacher autonomy. The teachers' way of framing the question suggested that what the parents perceived as a vice was, in fact, a virtue.)

The data also indicated that dialogue and discussion could normally lead those who initially viewed educational questions from different frames of reference to see the issue from their opponent's perspective. Contrary to what was anticipated, however, increased understanding did not normally lead participants--in this case primarily teachers and the principal--to alter their initial position. This was especially surprising since parents were asking for very little. Parents, even those who the school staff perceived as critics of the school, were more than willing to grant teachers a large amount of autonomy (possibly because parents realized that they had very little choice in the matter) and to support most of what went on in the school. Parents asked only that procedures be developed to insure the

existence of a minimal amount of continuity across classes on the same grade level and between different grade levels in a minimal number of fundamental areas so that students would not be at a disadvantage as they moved from class to class and school to school. School personnel refused to honor even this modest request.

The question which confronted us during data analysis was why. As indicated above, the evidence clearly demonstrates that the problem was not that school people could not adequately understand the parents' position. On the contrary, school people could frequently articulate the parents' position more clearly than the parents could themselves. While the evidence can clearly rule out the lack-of-understanding interpretation, however, it does not unequivocally rule <u>in</u> any single interpretation. The final report, in fact, presents multiple interpretations.

One of these which is explicated only briefly in the final report is a quasi-psychological account which is based on evidence suggesting that teachers never accepted the assumption that non-professionals could teach professionals anything. This quasi-psychological interpretation is somewhat similar to a more political interpretation suggested by Miles (1983) in his critique of a paper based on the final report. Miles indicated that the school in which the evaluation took place was similar to a school studied by Gold and himself (1981) and that like the data in their study the data concerning the deliberative evaluation could be seen as a manifestation of "a basic tension... between professional judgement and citizen participation,...

between expertise and democracy as a basis for decision-making" (p.5).

Another interpretation discussed briefly in the final report emphasized the fact that the school historically had been radically different from all other schools in the district and suggested that because of this historical difference, the school's staff had consistently been forced into a defensive posture which now makes them resistant to virtually any proposal for change made by an outsider. Teachers, themselves, after reading the final report, proposed yet another historically based interpretation. They indicated that the most vocal parents on the evaluation team had in the past frequently made what the teachers perceived as unreasonable, self-centered demands, and the teachers indicated that this history explained the staff's lack of receptivity to parent requests.

Although I believe each of the four explanations mentioned above is plausible -- I would probably go so far as to even say valid -- I also would not alter the position I took in the study's final report:

> The most powerful explanations... are functional arguments relating to the internal workings of the school... [These arguments] challenge the conceptual juxtaposition of (1) conflict resulting from individuals attempting to maximize their personal preferences and (2) conflict resulting from individuals viewing the same situation from different perspectives. The case being discussed here clearly demonstrates that protecting one's perspective can be just as much a self interest as protecting one's salary and can result in the use of tactics which are every bit as political as those used by, say, a teachers' union negotiating for higher salaries.
>
> At least in the context where the evaluation was conducted, the school's philosophy served to do

> much more than merely articulate "facts" and values and integrate the two. The school's philosophy functioned more as a theology than as a scientific theory which is modifiable by evidence and argument. At Kenton Lane Elementary School words such as creativity and individualization seemed to form something akin to a sacred language and [certain unique curriculum and instruction] procedures... appeared to function as a kind of religious ritual. As is often the case with religion, what is practiced is not entirely isomorphic with what is professed. But religious beliefs often serve to generate emotional commitment and a sense of unity within a religious community, nevertheless, and Kenton Lane's philosophy clearly served both these functions. (Donmoyer, 1983, p. 30).

To establish the validity of this interpretation, the report focuses upon two contradictions which are apparent in the school: (1) the school's staff thought of the school as "uniquely different," yet the staff confronted many of the same organizational realities (e.g., teaching a group of students rather than tutoring an individual child) and many of the same constraints (e.g., preparing students for future levels of schooling and for a society that does not aspire to be "uniquely different") which all teachers in public schools confront; and (2) talk of the school as "uniquely different" implied some sort of unified program, yet the school was made up of a collection of strong individualists who often approach teaching in decidedly different ways.

The report attempts to document the fact that these contradictions exist, that the school's vaguely defined philosophy functions to "resolve" these contradictions for staff members, and that acceptance of parents' requests and logic would have undermined a philosophy which, for staff

members, was highly functional. The documentation is extensive, and most who have read the final report -- including those who have proposed alternative explanations -- have found the case convincing.

The problem is that I probably could have constructed convincing cases for each of the four other interpretations mentioned above. While the empirical evidence is adequate for ruling out possible interpretations (e.g.,the interpretation that failure to accommodate parents' concerns resulted from inadequate communication) it does not provide an adequate basis for choosing the interpretation which is most valid. This inadquacy results, in part, from the fact that social phenomena are complex and have multiple interacting causes (Cronbach, 1975) and in part from the fact that we have multiple languages with which to talk about social phenomena. In the study being discussed here, for example, political scientists would probably prefer Gold and Miles' (1981) language, while emic-oriented anthropologists would undoubtedly choose participants' own characterization of events as their own. Only certain types of sociologists and anthropologists would likely be partial to my functionalist account.

Given this state of affairs, was my choice of a functional explanation arbitrary and capricious? More generally, whenever researchers choose between two or more empirically plausible interpretations, are their choices inevitably non-rational? Or, to put the matter another way, can we no longer accommodate what Phillips calls "the concept of a mistake" when we must

choose the best among a group of empirically adequate interpretations?

The Solution

Toulmin's analysis, I believe, clearly indicates that the answer to each of the above questions is no. Political scientists and emic-oriented anthropologists will likely prefer certain interpretations over others because certain interpretations are more relevant to the particular purposes to which their disciplines or subdisciplines are committed. Their choices are far from arbitrary, however.

Similarly, my decision to focus upon a functionalist account of the phenomena was based on the fact that such an interpretation was most relevant to the particular purposes which undergirded my study. These purposes were undoubtedly more idiosyncratic and less discipline-bound than political scientists'. (At best, Toulmin would probably say that researchers who share an interest in the notion of deliberation and a desire to explore the applicability of this notion for educational decision making (e.g. Reid, 1978; Schwab, 1969, 1971, 1973, 1983) constitute a quasi-discipline). The more personal nature of my purpose, however, in no way effects the rationality of the choice I made. The purpose of my study was to explore the viability of the notion of deliberation in the context of program evaluation. More specifically the study was a sort of informal test of the assumptions on which a deliberative approach to evaluation rested. Of all the interpretations which appeared empirically viable, the functional explanation

addressed these concerns most directly. (The data, in fact, directly challenged a fundamental assumption on which a deliberative approach to evaluation rested.) The decision to focus upon this interpretation, therefore, was highly rational and certainly not arbitrary.

To summarize, questions of validity ultimately require that two subquestions be asked and answered: (1) What are the facts? and (2) What is our purpose? The answer to the first question helps eliminate implausible interpretations. The answer to the second provides the basis for choosing from among empirically plausible alternatives.

Further Implications

In addition to resolving a specific data analysis problem, I believe the above discussion can also bring a modicum of clarity to the often confused and increasingly tiresome discussion and debate about two general research orientations which, for want of better terminology, we have named (or, to be more accurate, misnamed) qualitative and quantitative. Therefore, I will conclude this paper with three, relatively succinct statements about "qualitative" and "quantitative" research which I believe follow from the preceding analysis.

(1) While legitimate differences exist between the research orientations we have dubbed qualitative and quantitative, this distinction has more to do with the nature of the questions addressed and the personal

preferences of researchers and researchers' audiences than with fundamental differences in epistemology. The logic of validation is the same in both kinds of research. (Indeed, it is constant for all forms of human inquiry.) Most decisions researchers must make (e.g., what procedures to employ, what should count as evidence) are normally most appropriately characterized as strategic in nature (i.e., decisions based on an assessment of which choice will best accomplish the researcher's particular purposes) rather than as decisions based on epistemological considerations.

(2) To date the debate between advocates of qualitative and quantitative orientations has often been non-rational and, therefore, the issues that have been debated have remained largely unresolved and unresolvable. The only resolution that has been viable has been quantitative researchers' token acceptance of the least radical versions of qualitative research, i.e., versions whose purposes are consistent with the purposes undergirding more traditional methods of inquiry. More radical versions of qualitative research (e.g. Eisner's (1979) educational criticism) which are oriented toward different purposes (Donmoyer, 1981) are either dismissed as unacceptable or ignored. There has been virtually no serious debate over whether non-traditional purposes are worth pursuing, and assumptions about whether specific procedures can, in fact, accomplish specific non-traditional purposes have seldom been critically examined. This absence of rational debate results, in part, from the fact that many traditional, quantitatively oriented researchers have either become <u>de facto</u> relativists or remain true believers in a positivist epistemology which has long been discredited and

which, despite Phillips' recent salvage attempt, remains an indefensible theory of knowledge.

(3) Erroneous epistemological positions have also been put forth by advocates of qualitative research, and these positions have also served to inhibit rational discourse. For example, the analysis in this paper suggests that qualitative researchers' talk of findings emerging from the data is no more defensible than the naive empiricism of quantitative researchers. Rational discourse is only possible when notions such as grounded theory and holism are defined and defended in strategic rather than epistemological or ontological terms.

Just as validity questions can only ultimately be resolved by considering the researcher's purpose, so a consideration of purpose is central to the rational resolution of all significant research disputes. Ultimately all decisions a researcher makes should hinge not on the nature of the phenomena being studied nor on some sort of litmus test of the empirical purity of evidence. Rational judgements are ultimately grounded in the purposes the researcher has set.

References

Conklin, H. "Hanunoo Color Categories". Southwestern Journal of Anthropology, 11:339-344.

Dewey, J. Democracy and Education. New York: The Macmillan Company, 1916.

Donmoyer, R. Evaluation as Deliberation: Issues, Assumptions, and Findings from an Exploratory Study. Final Report (Grant No. NIE-G-81-0083) submitted to the National Institute of Education, April, 1983.

Eisner, E. The Educational Imagination. New York: Macmillan, 1979.

Feyerabend, P. "Against Method: Outline of an Anarchistic Theory of Knowledge." In M. Rodner and S. Winokur (Eds.), Analyses of Theories and Methods of Physics and Psychology. Minneapolis: University of Minnesota Press, 1970.

Gold, B. and Miles, M. Whose School Is It, Anyway: Parent Teacher Conflict Over an Innovative School. New York: Praeger, 1981.

Hanfling, O. Logical Positivism. Oxford: Basil Blackwell, 1981.

Kuhn, T. The Structure of Scientific Revolutions. Chicago: University of Chicago Press, 1962.

Miles, M., Critique of Session on Problems and Promises of Collaborative Research, 1983 Annual Meeting of The American Educational Research Association, Montreal.

Nagel, E. The Structure of Science. New York: Harcourt, Brace and World, 1961.

Phillips, D. "After the Wake: Postpositivistic Educational Thought". Educational Researcher , 12: 4-12.

Reid, W. Thinking about the Curriculum. London: Routledge and Kegan Paul, 1978.

Schwab, J. "The Practical: A Language for Curriculum" School Review. (Nov., 1969), 76,1.

_____. "The Practical: Arts of the Eclectic," School Review. (1971),79,4.

_____. "The Practical 3: Translation into Curriculum" School Review, (1973), 81,4.

_____. "The Practical 4: Something for Curriculum Professors to Do," Curriculum Inquiry (1983) 13,3.

Smith, Land Geoffrey W. The Complexities of an Urban Classroom. New York: Holt, Rinehard, and Winston, 1968.

Toulmin, S. Foresight and Understanding. New York: Harper and Row, 1961.

_____. Human Understanding. Princeton: Princeton University Press, 1972.